

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 16, 2001, 18:14:46 ; Search time 45.43 Seconds  
(without alignments)  
1872.598 Million cell updates/sec

Title: US-09-405-504A-53  
Perfect score: 3384

Sequence: 1 MLLGASLVGLLFSKLVKL.....RYVPLDQAYSIQAGEEKL 643

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_16:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_unclassified:\*
- 13: sp\_vertebrate:\*
- 14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3152	93.1	641	4	O95186
2	2433	71.9	506	11	O88562
3	1427.5	42.2	671	5	O9W1V9
4	1408.5	41.6	690	5	O9VKU1
5	1292.5	38.2	661	5	O9W185
6	1292	38.2	650	5	O19878
7	1281	37.9	655	5	O18916
8	1122	33.2	619	4	O9Y2P4
9	1053.5	31.1	614	11	O88561
10	1021	30.2	620	11	O70550
11	984.5	29.1	690	4	O9Y2P4
12	969.5	28.6	690	11	O9ES38
13	960	28.4	662	11	O88694
14	918.5	27.1	608	2	O9HZV4
15	886.5	26.2	597	2	O05307
16	877.5	25.9	623	11	O88560
17	801	23.7	643	3	O42633
18	742	21.9	669	3	O60021
19	408	12.1	502	2	O53551

20	389	11.5	503	2	O53306
21	375	11.1	569	1	O28502
22	371.5	11.0	555	2	O85737
23	364.5	10.8	513	2	O07610
24	358	10.6	532	2	P72007
25	346.5	10.2	535	2	O9KHL1
26	332	9.8	599	1	O9Y9K4
27	322.5	9.5	560	2	P95227
28	318.5	9.4	504	2	O9F7P5
29	318	9.4	661	3	O13441
30	311.5	9.2	549	2	O31826
31	310	9.2	2723	2	O30479
32	301.5	8.9	486	2	O34837
33	300.5	8.9	549	2	O9R915
34	298.5	8.8	593	1	O29233
35	297	8.8	545	10	O9SWH8
36	296.5	8.8	511	2	O9KZC1
37	294.5	8.7	552	1	O28762
38	294	8.7	542	10	O9FOY7
39	294	8.7	554	2	O07411
40	293.5	8.7	512	3	O74976
41	292	8.6	584	2	O9RTR4
42	291	8.6	444	2	O9X4W6
43	291	8.5	535	10	O81139
44	286	8.5	8563	2	O54297
45	285.5	8.4	517	2	O9JZ62

## ALIGNMENTS

RESULT 1

O95186					
ID	O95186	PRELIMINARY;	PRT;	641	AA.
AC	O95186;				
DT	01-MAY-1999 (TrEMBLrel. 10, Created)				
DT	01-MAY-1999 (TrEMBLrel. 10, Last sequence update)				
DT	01-MAY-2000 (TrEMBLrel. 13, Last annotation update)				
DE	FATTY ACID TRANSPORT PROTEIN.				
GN	FATP.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
OC	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=HEART;				
RX	MEDLINE=99096471; PubMed=9878842;				
RA	Fitscher B.A., Riedel H.D., Young K.C., Stremmel W.;				
RT	"Tissue distribution and cDNA cloning of a human fatty acid transport protein (hFATP4).";				
RL	Biochim. Biophys. Acta 1443:381-385(1998).				
DR	EMBL; AF055899; AAD11623.1; -				
DR	InterPro; IPR000566; -				
DR	InterPro; IPR000873; -				
DR	Pfam; PF00501; AMP-binding; 1.				
DR	PROSITE; PS00455; AMP_BINDING; 1.				
DR	PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.				
SQ	SEQUENCE 641 AA; 71431 MW; 8C24F76C9BF81378 CRC64;				

Query Match 93.1%; Score 3152; DB 4; Length 641;  
Best Local Similarity 94.7%; Pred. No. 6.6e-256;  
Matches 610; Conservative 4; Mismatches 26; Indels 4; Gaps 3;

OY	1	MLLGASLVGLLFSKLVKLPTQVGSLSFLYLXGSGWRFIRVFIKTRRDFGGLVLL 60	
Db	1	MLLGASLVGLLFSKLVKLPTQVGSLSFLYLXGSGWRFIRVFIKTRRDFGGLVLL 60	
OY	61	KVAKVRQCLQERTVPILPASTVRRHPDKTALIFGTDHTWTFRLDEYSSSVANFLQA 120	
Db	61	KVAKVRQCLQERTVPILPASTVRRHPDKTALIFGTDHTWTFRLDEYSSSVANFLQA 120	

QY 121 RGLASGDVAAIFMENEFVGLWGLMAKLGVEAALINTNLRDALLHCLTTSRARALVFG 180  
DB 121 RGLASGDVAAIFMENEFVGLWGLMAKLGVEAALINTNLRDALLHCLTTSRARALVFG 180  
QY 181 SEMASAICEVHASLDSLSLFCSGSWEPCGAVPPSTHLDPLKAPKHLPCSPDGFDTK 240  
DB 181 SEMASAICEVHASLDSLSLFCSGSWEPCGAVPPSTHLDPLKAPKHLPCSPDGFDTK 240  
QY 241 LFYIYSGTGLPKAAIVHVSRYRMAALVYGFMRPNNDIVYDCLPLYHSAGNIVGIGQ 300  
DB 241 LFYIYSGTGLPKAAIVHVSRYRMAALVYGFMRPNNDIVYDCLPLYHSAGNIVGIGQ 300  
QY 301 CLLHGMTVTVIRKFSASRFDWDCIKYNCITVQYIGELCRYLLNQPPREAEHOQVRMALG 360  
DB 301 CLLHGMTVTVIRKFSASRFDWDCIKYNCITVQYIGELCRYLLNQPPREAEHOQVRMALG 360  
QY 361 NGLRQSIWTFNSRPHIPOVAEYFAGTECNSLGNFDSOVGACGNSRILSFVYPIRLVR 420  
DB 361 NASGSPSGPTFPAASTYPRWLSSTG-PECNCSLGNFDSOVGACGNSRILSFVYPIRLVR 419  
QY 421 VNEDTMELIRGPGVCIPOCPGPGOLVGRITOKDPLRRFDGVLNQGANNKIAKDVFKK 480  
DB 420 VNEDTMELIRGPGVCIPOCPGPGOLVGRITOKDPLRRFDGVLNQGANNKIAKDVFKK 479  
QY 481 GDQAYLTGDLVMDLGYLYFRDRTGDTFRWKGNSVSTTEVEGTLRLLDMADVAVGYE 540  
DB 480 GDQAYLTGDLVMDLGYLYFRDRTGDTFRWKGNSVSTTEVEGTLRLLDMADVAVGYE 539  
QY 541 VPGTEGRAGMAAVASPTGNCDLERFAQVLEKELPLYPARIFLRLPELHKTGYFKQTE 600  
DB 540 VPGTEGRAGMAAVASPTGNCDLERFAQVLEKELPLYPARIFLRLPELHKTGYFKQTE 599  
QY 601 LRKEGEPDPAIVKPL-FYLDQAOKGRVPLDQEAYSRIQAGEEKL 643  
DB 600 LRKEAFDPAIVKRCISIYI--EKGRIYPLDQEAYSRIQAGEEKL 641

RESULT 2  
O88562 PRELIMINARY; PRT; 506 AA.  
AC O88562;  
DT 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE FATTY ACID TRANSPORT PROTEIN 4 (FATP4) (LONG-CHAIN FATTY ACID  
TRANSPORT PROTEIN 4) (FRAGMENT).  
GN SLIC27A4.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98337965; PubMed=9671728;  
RA Hirsch D., Stahl A., Lodish H.F.;  
RT "A family of fatty acid transporters conserved from mycobacterium to  
man.";  
RL Proc. Natl. Acad. Sci. U.S.A. 95:8625-8629(1998).  
CC -I- FUNCTION: INVOLVED IN TRANSLOCATION OF LONG-CHAIN FATTY ACIDS  
ACROSS THE PLASMA MEMBRANE. MAY PLAY A PIVOTAL ROLE IN REGULATING  
AVAILABLE LONG-CHAIN FATTY ACID SUBSTRATES FROM EXOGENOUS SOURCES  
IN TISSUES UNDERGOING HIGH LEVELS OF BETA-OXIDATION OR  
TRIGLYCERIDE SYNTHESIS.  
CC -I- TISSUE SPECIFICITY: HEART, BRAIN, LUNG, LIVER, AND KIDNEY.  
CC -I- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT  
COVALENT BINDING OF AMP TO THEIR SUBSTRATE.  
CC EMBL; AF027259; AAC40188.1; -;  
DR MGD; MGI:1347347; SLIC27a4.  
DR InterPro: IPR000873; -;  
DR Pfam: PF00501; AMP-binding; 1.  
DR PRINTS: PR00154; AMPBINDING  
DR PROSITE: PS00455; AMP\_BINDING; 1.  
KW Glycoprotein; Lipid transport; Transport.

FT NON\_TER 1 1  
FT CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 253 253 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 378 378 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 506 AA; 56902 MW; 4E8BC0C2173326D CRC64;  
Query Match 71.9%; Score 2433; DB 11; Length 506;  
Best Local Similarity 91.2%; Pred No. 1.1e-195;  
Matches 455; Conservative 23; Mismatches 21; Indels 0; Gaps 0;  
QY 145 GMAKLGVEAALINTNLRDALLHCLTTSRARALVFGSEMAAICEVHASLDSLSLFCSG 204  
DB 8 GMAKLGVEAALINTNLRDALLHCLTTSRARALVFGSEMAAICEVHASLDSLSLFCSG 67  
QY 205 SWEGCAGVPPSTHLDPLKAPKHLPCSPDGFDTKLFYIYSGTGLPKAAIVHVSRY 264  
DB 68 SWEPSTVPVSTHLDPLLEDAKPHLPDHPKGFDTKLFYIYSGTGLPKAAIVHVSRY 127  
QY 265 RMAALVYGFMRPNNDIVYDCLPLYHSAGNIVGIGQCLLHGMTVTVIRKFSASRFDWDCI 324  
DB 128 RMAALVYGFMRPNNDIVYDCLPLYHSAGNIVGIGQCLLHGMTVTVIRKFSASRFDWDCI 187  
QY 325 KYNCTIYOYIGELCRYLLNQPPREAEHOQVRMALGNSLQSIWTFNSRPHIPOVAEY 384  
DB 188 KYNCTVQYIGELCRYLLNQPPREAEHRKVRMALGNSLQSIWTFNSRPHIPOVAEY 247  
QY 385 GATECNSLGNFDSOVGACGNSRILSFVYPIRLVRVNEEDTMELIRGPGVCIPOCPGEP 444  
DB 248 GATECNSLGNFDSOVGACGNSRILSFVYPIRLVRVNEEDTMELIRGPGVCIPOCPGEP 307  
QY 445 GQLVRIIQQDPLRRFDGVLNQGANNKIAKDVFKEGQDQAYLTGDLVMDLGYLYFRDR 504  
DB 308 GQLVRIIQQDPLRRFDGVLNQGANNKIAKDVFKEGQDQAYLTGDLVMDLGYLYFRDR 367  
QY 505 TGDTRFKWGENVSTTEVEGTLRLLDMADVAVGYEVPVTEGRAGMAAVASPTGNCDLER 564  
DB 368 TGDTRFKWGENVSTTEVEGTLRLLDMADVAVGYEVPVTEGRAGMAAVASPTGNCDLER 427  
QY 565 FAQVLEKELPLYPARIFLRLPELHKTGYFKQTELRKEGDFDPAIVKDPFLYLDQAOKGR 624  
DB 428 FAQVLEKELPLYPARIFLRLPELHKTGYFKQTELRKEGDFDPAIVKDPFLYLDQAOKGR 487  
QY 625 YVPLDQEAYSRIQAGEEKL 643  
DB 488 YVPLDQEAYSRIQAGEEKL 506

RESULT 3  
Q9W1V9 PRELIMINARY; PRT; 671 AA.  
AC Q9W1V9;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE CG3037 PROTEIN.  
GN CG3037.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,

[illegible]



QY 386 ATECNSLGNFDSQVACGNSRILSFVYPIRLVVRVNEDETMELIRGDCVCPQGPBG 445  
 DB 405 ATEGNSLINTNRVGAIGFVYVYSSLPVQVLRDEYTGELLKDSKGCRCQPGQAG 464  
 QY 446 OLVGRIQKDLRRFGYLNQGANKKIAKDFKGDQAYLGTGDLVMDLGLYFRDRRT 505  
 DB 465 LVLGVKVDARRAVAFHGYADKGASEQKLLRNVTSGDFFNSGDMVVRDILGYFFKORT 524  
 QY 506 GDFRKGNGVSTVEGTLRLDMDADVAVYGVVEPTGEGRAGMAVAPSGNCDLERF 565  
 DB 525 GDFRKGNGVSTVEGTLRLDMDADVAVYGVVEPTGEGRAGMAVAPSGNCDLERF 565  
 QY 566 AQLERKELPLYPARLRLPLHKLTKYFKOKTELKKEGPPDAIVKDFLYLDAQKGRY 625  
 DB 585 SVVLRSGLPPYARLRLPLHKLTKYFKOKTELKKEGPPDAIVKDFLYLDAQKGRY 625  
 QY 626 VPLDQAYSRIOAGEKL 643  
 DB 644 RPLSQHQHLLRSGKAGL 661

RESULT 6  
 Q19878 ID Q19878 PRELIMINARY; PRT; 650 AA.  
 AC Q19878;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE FATTY ACID TRANSPORTER PROTEIN A (F28D1.9 PROTEIN).  
 GN F28D1.9.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Baynes C.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: INVOLVED IN TRANSLLOCATION OF LONG-CHAIN FATTY ACIDS (BY  
 CC SIMILARITY).  
 CC -!- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT  
 CC COVALENT BINDING OF AMP TO THEIR SUBSTRATE.  
 DR EMBL; 270684; CAA94602.1;  
 DR WormPep; F28D1.9; CE05749.  
 DR InterPro; IPR000873;  
 DR Pfam; PF00501; AMP-binding; 1.  
 DR PROSITE; PS00455; AMP-BINDING; 1.  
 KW Glycoprotein; Lipid transport; Transmembrane; Transport.  
 FT TRANSMEM 23 43 POTENTIAL.  
 FT TRANSMEM 137 157 POTENTIAL.  
 FT TRANSMEM 286 306 POTENTIAL.  
 FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 518 518 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 650 AA; 72759 MW; F2D16649EAL3288A CRC64;

Query Match 38.28; Score 1292; DB 5; Length 650;  
 Best Local Similarity 41.8%; Pred. No. 9.7e-100;  
 Matches 272; Conservative 123; Mismatches 228; Indels 28; Gaps 10;  
 QY 7 LVGVLLFSLKVL--KLPTQVGSLLFLYLGSGGWRFRIVFIKTRRDFGLVLLKVA 64  
 DB 6 LVTVMLLTAVIAQNLPGVILAGVILYITVWHGDFIYRSYLTNRDLTGLALIEVKI 65  
 QY 65 VKROCLQRRVTPILFASTVRRHPDKTALIFEGDTHWTFQQLDEYSSVANFLOARGLA 124  
 DB 66 DLWRLHONKHGHELFDLIVKNNKPNAMIDTETTYAEFNAHCNRVANYFQGLGRY 125  
 QY 125 SGDVAIFMENRNFVGLWGLMAKLGVEAALINTNLRDALLHCLTTSRARALVFGSEMA 184  
 DB 126 SGDVAALYMENRNFVGLWGLMAKLGVEAALINTNLRDALLHCLTTSRARALVFGSEMA 185

QY 185 SAICEVHASLDPSL-----SLFCSGSWEP---GAVPSTPEHLDPKAPKHLPSCPDK 235  
 DB 186 NIMLD---AIDQKLFDEVEGIEVSVG--EPKNSGFKNLKKLDAQTITTEPKTLDIV--- 237  
 QY 236 GFTDKLFYITSGTGLPKAAIIVHVSRYRMAALVYVYGFMRPNDIVYDCLPLVHSAGNI 295  
 DB 238 DFKSLICFIYSGTGMKAAVYVYGFMRPNDIVYDCLPLVHSAGNI 297  
 QY 296 VIGOCCLHGMTVVRKFKSASREFWDDCIKYNCTIVYIGELCYLLNQPPREAEHQV 355  
 DB 298 LGVQALLGGSCVIRKKFSASNEWDCVKYDCTVSYIGEICRYLLAQPWVEESHRM 357  
 QY 356 RMALNGLRQSIWTFNFSRPHIPOVAEYFATGECNCSLGNPDSQVACGF--NSRILSFV 413  
 DB 358 RLLVNGLRARAEIWPFDVDRFRV--RIGELYGSTEGTSSLVNIDHGVGACGFLPSPLTKM 416  
 QY 414 YPIRLVRVNEDETMELIRGDCVCPQGPGLVIRGDRGQVGLVRIKQDLRLRDEGVLNQCANNKI 473  
 DB 417 HPVRLIKVDDVTGGAIRTSGLCIACNPGSGAMVSTRKNNPLLQFEGYLNKKTNNKI 476  
 QY 474 AKDFKKGDDQAYLGTGDLVMDLGLYFRDRRTGDTFRWKGNGVSTTEVEGTLRLDMDAD 533  
 DB 477 IRDVFAGKGDSCFLTGDLHWDRLGVYVFKORTGDTFRWKGNGVSTTEVEAILHPITGLSD 536  
 QY 534 VAVYGVVEPTGEGRAGMAVAPSGNCD-----LERTAQVLEKELPLYPARLRLPEL 588  
 DB 537 ATVYGVVEPTGEGRAGMAVAPSGNCD-----LERTAQVLEKELPLYPARLRLPEL 588  
 QY 589 HKTGYFQKTELKKEGFPDAIVKDFLYLDAQKGRVYVPLDQAYSRIOAG 639  
 DB 597 EKTGTFKLKTNLQRLGIMDA--PDSYIYINSENRENVFPFNDLRCKVSLG 646

RESULT 7  
 Q18916 ID Q18916 PRELIMINARY; PRT; 655 AA.  
 AC Q18916;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE FATTY ACID TRANSPORTER PROTEIN B.  
 GN D1009.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Favello T.;  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Waterston R.;  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: INVOLVED IN TRANSLLOCATION OF LONG-CHAIN FATTY ACIDS (BY  
 CC SIMILARITY).  
 CC -!- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT  
 CC COVALENT BINDING OF AMP TO THEIR SUBSTRATE.  
 DR EMBL; 040938; AAA81698.1;  
 DR HSSP; P14687; IAMU.  
 DR WormPep; D1009.1; CE04285.  
 DR InterPro; IPR000873;  
 DR Pfam; PF00501; AMP-binding; 1.  
 DR PROSITE; PS00455; AMP-BINDING; 1.  
 KW Glycoprotein; Lipid transport; Transmembrane; Transport.  
 FT TRANSMEM 10 30 POTENTIAL.  
 FT TRANSMEM 142 162 POTENTIAL.  
 FT TRANSMEM 295 315 POTENTIAL.  
 FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 282 282 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 522 522 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 655 AA; 73514 MW; 2AB3BF7B198D14C CRC64;

Query Match 37.9%; Score 1281; DB 5; Length 655;  
 Best Local Similarity 41.4%; Pred. No. 8.2e-99;  
 Matches 266; Conservative 116; Mismatches 230; Indels 30; Gaps 8;

QY 23 TOVGFSLFLYLGSGGWRFT-----BVFIKTRRRDIFGGLVLKAKVR 67  
 DB 14 TEVYAVLVNVSWEKFFVIGYVVRLLRTDGRALATLPRDFAGLKLLISVKSTIR 73

QY 68 QCLQERTVILFASIVRRHPDKTALIFEGDTHWTFRQLDEYSSVANFLQARGLASGD 127  
 DB 74 GLFKKORPIEHLINQVQKHPVAIEIESGRLTYQELNALANQYALNLYSEGYKMGD 133

QY 128 VAAIFMENRNEFVGLWLGMAKLGVEAALINTNLRDALLHCLTTSRARALVFGSE---MA 184  
 DB 134 VVALFENSIDFAIWLGLSKIGVVSNAFINSNKLPLELAHSINVSCKSCIINILLPMF 193

QY 185 SATCEVHASLDPSLSFCSSWEPGAVPSTEHLDPLLDKADPKHLPSCDPKGTOKLFIY 244  
 DB 194 KAAREKNLISD-EIHVFLAGTQVDRHSRSLQDLHLFSEDEP---EVIDGLNFRSVLCYI 249

QY 245 YTSCTTGLPKAAIIVHSRYRMAALVYVGRMRPNDIVDCLPLYHSAGNIVGICGLLH 304  
 DB 250 YTSCTTGNPAPAVIKHFRYFIAMGAKAPGINKSDVYITMFWYHSAAGIMGIGSLIAF 309

QY 305 GMTVVIRKFSASREWDDCIYKNTCTIVQYIGELCYRLLNQPPREAEHQVIRMALGNGLR 364  
 DB 310 GSTAVIRKFSASFNWKCVMKYNVTATQYIGETICRYLLAANPCPEKQHNVRMLWNGNLR 369

QY 365 OSIWNFSRPHIPOVAEFGATECNSLGNFDSQVAGCGFNSRI--LSFVPIRLVRYN 422  
 DB 370 GQIWKEVGFGRGKIGELGSGTEGNSVNVNHNHVGACGFMPYIPHIGSLPVRLIKVD 429

QY 423 EDTMELIRGPGVCIPQCPGEGQLYGRITQKDLRRFDGLYLNQAGNNKIAKDVFKGD 482  
 DB 430 RATGELERDNGLCVPCVPGETGEMGVVKEIKDILLKFGYVSEGDATKKIYRDVFKHD 489

QY 483 QAYLTGVLVMDLGYLYIFRDRGDTFRWKGENSVTEVEGTLRLLDMADVAVYGVVFP 542  
 DB 490 KVFASGDLHWDLDGLYLYFYFDRGDTFRWKGENSVTEVEGTLRLLDMADVAVYGVVFP 549

QY 543 GTEGRAGMAVASPTGNCLELRE---AOVLEKELPLYARPIRLPLPELHKTGTGKFK 598  
 DB 550 KMEGRAGMAGIVVDGT-DVEKFIADTISRLTENLASYAIPVIRLCKEVDRTGTFKLRK 608

QY 599 TELRKEGDFPAIVK-DPLFYLDQAGRYVPLDQAYSRIQAG 639  
 DB 609 TDLQKQGYDLVACKGDPFIYVWSAAEKSYKPLTKMQQDIDTG 650

RESULT 8  
 QY2P4 PRELIMINARY; PRT; 619 AA.

ID AC QY2P4;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE VERY LONG-CHAIN ACYL-COA SYNTHETASE HOMOLOG 1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Steinberg S.J., Watkins P.A.;  
 RT "Human Very Long-Chain Acyl-CoA Synthetase Homolog 1.";  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF064254; AD29443.1;  
 DR InterPro: IPR000873;  
 DR Pfam: PF00501; AMP-binding; 1.  
 DR PROSITE: PS00455; AMP\_BINDING; 1.  
 SQ SEQUENCE 619 AA; 70111 MW; 20264CE3FBB44FF5 CRC64;

Query Match 33.2%; Score 1122; DB 4; Length 619;  
 Best Local Similarity 39.2%; Pred. No. 1.7e-85;  
 Matches 258; Conservative 107; Mismatches 230; Indels 64; Gaps 12;

QY 2 LIGASLVGLFLSKVLKLPWTVQVGSLLFLYLGSGGWRFTIRVFIKTRRDIRFGGLVLLK 61  
 DB 8 VLGAGW-VLHFLQ-----KLLFPYFWDDEFVLKV-----VLII 41

QY 62 VKAKVRQCLQERRIVPILFASIVRRHPDKTALIFEGDTHWTFRQLDEYSSVAN-FLQA 120  
 DB 42 IRLKYEKRGELVLDKFLSHAKRQPKPKPIIYEG--DIYTDQVDKRSRVAHVNLH 99

QY 121 RGLASGVAAIFMENRNEFVGLWLGMAKLGVEAALINTNLRDALLHCLTTSRARALVFG 180  
 DB 100 SSLKGGDVALMSNEPDEFVHWFGLAKLGCVAFATNIRNSNLLNCIRACGPRALVVG 159

QY 181 SEMASICEVHASLDPSLSFCSG-----SWEPGAVPSTEHLDPLLDKADP 226  
 DB 160 ADLLGTVEEILFSLSENISVWGMKDSVPQGVISLKEKLSSTSPDEPVRSHHVSLKLS-- 217

QY 227 KHLPSCKDKGFTDKLFIYTSCTTGLPKAAIIVHSRYRMAALVYVGRMRPNDIVDCL 286  
 DB 218 ----TC-----LYFTSGTGLPKAAVISQLVLRGSA-VLWAFGCTAHDIVITL 263

QY 287 PLYHSAGNIVGICGLHGMFTVIRKFSASREWDDCIYKNTCTIVQYIGELCYRLLNQPP 346  
 DB 264 PLYHSSAAILGISGVELGATCVLKKKFSASQFWSQCKYDVTVFQYIGELCYRVLCKSK 323

QY 347 REAENQHOVRMALGNGLRQSLTWNFSRPHIPOVAEFGATECNSLGNFDSQVAGCGFN 406  
 DB 324 REGEKDHVRALGNGHRSVDVREFLGRFNGIKVCELYAATESISPMNYTGRIGART 383

QY 407 SRILSFVYPIRLVRVNEDELMELIRGPGVCIPQCPGEGQLYGRITQKDLRRFDGLYLNQ 466  
 DB 384 NLFYKLLSTFDLIKYDFQKDEPMRNEQGWCIHVKKGEPLLSRVNAKNP---FFGYAGP 440

QY 467 GANNK-KIAQDVFKKGOAYLTGDLVMDLGYLYIFRDRGDTFRWKGENSVTEVEGTL 525  
 DB 441 YKHTKDKLLCDLVFKKGDVYLTGDLIVDQDNFLYFDWRTGDTFRWKGENVATTEADVI 500

QY 526 SRLIDMAADVAVYGVVPCTEGRAGMAA-VASPTGNCLELREFAOVLKELPLYARPIRL 584  
 DB 501 GMLDEIQANVYGAISGYEGRAGMASILKAPNTSLDERKVEYQVVTFLPAYACPRLRI 560

QY 585 LPELHKTGTYKFKTELKKEGDFPAIVKDPFLFYLDQAGRYVPLDQAYSRIQAGEKL 643  
 DB 561 QEKMEATGCTKLLKHQLVEDGFPNPKISEPLYFMDNLKKSYYLLTRELIDQIMLGEIKL 619

RESULT 9  
 O88561 PRELIMINARY; PRT; 614 AA.

ID AC O88561;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE FATTY ACID TRANSPORT PROTEIN 3 (FATP3) (LONG-CHAIN FATTY ACID  
 TRANSPORT PROTEIN 3) (FRAGMENT).  
 GN SLC27A3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=98337965; PubMed=9671728;  
 RA Hirsch D., Stahl A., Lodish H.F.;  
 RT "A family of fatty acid transporters conserved from mycobacterium to  
 man.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 95:8625-8629(1998).  
 CC -!- FUNCTION: INVOLVED IN TRANSLOCATION OF LONG-CHAIN FATTY ACIDS

CC ACROSS THE PLASMA MEMBRANE. MAY PLAY A PIVOTAL ROLE IN REGULATING  
CC AVAILABLE LONG-CHAIN FATTY ACID SUBSTRATES FROM EXOGENOUS SOURCES  
CC IN TISSUES UNDERGOING HIGH LEVELS OF BETA-OXIDATION OR  
CC TRIGLYCERIDE SYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. PLASMA MEMBRANE.  
CC -1- TISSUE SPECIFICITY: LUNG, LIVER, AND TESTIS.  
CC -1- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT  
CC COVALENT BINDING OF AMP TO THEIR SUBSTRATE.  
DR EMBL: AF072758; AAC40187.1; -  
DR MGD; MGI:1347358; SLC27a3.  
DR InterPro: IPR000873; -  
DR PRINTS; PF00501; AMP-binding; 1.  
DR PROSITE; PS00154; AMPBINDING.  
DR PROSITE; PS00455; AMP BINDING; 1.  
KW Glycoprotein; Lipid transport; Transmembrane; Transport.  
FT TRANSMEM 99 119 POTENTIAL.  
FT TRANSMEM 262 282 POTENTIAL.  
FT CARBOHYD 367 367 N-LINKED (GLCNAC...) (POTENTIAL).  
SQ SEQUENCE 614 AA; 67041 MW; 33C2A558CDD9D989 CRC64;

Query Match 31.1%; Score 1053.5; DB 11; Length 614;  
Best Local Similarity 39.9%; Pred. No. 9.4e-80;  
Matches 244; Conservative 85; Mismatches 219; Indels 63; Gaps 11;

QY 81 ASTVRRHPDKTALIFEGTDHTWFRQLDEYSSVAN-FIQARG----- 122  
DB 19 AYLAREQPTFLIHGA--QRSYAEARESNRIARAFRLARGTGGRRSGRGSTEEGA 76  
QY 123 -----LASGDVAAFMENENRPFVGLMGMAKLGVEAAALINTNLNRDA 164  
DB 77 RVAPPAGDAAARTGTAPPLAGCATVALLPAGDFLWTFGLAKAGLTATFVPTALRGP 136  
QY 165 LLHCLTTSARALYFGSEMAAICEVHASLDPSLSFCGSEWEPGAVPPST- 215  
DB 137 LLHCLSCGASALVATEFLE-----SLEPDLPALRAMGLHLWATPETNAGISNLL 189  
QY 216 -EHLDPDLLADKAPLPCSPDKGFTKLFVYVTSGLTPKAAIVVHSRYRMAALVYGF 274  
DB 190 SEAAQDVEPVGYL-SAP-QNIMDTCLYIFTSGLTPKAAISHLKVLCQCGF-YHLC 246  
QY 275 RMRPNDIVDCLPLYSAGNIVIGQCCLHGMVTVIRKFSASRFWDDCIKNTVQYI 334  
DB 247 GVHQEDVILALPLHYMSSLIGVGLGIGATVVLKPKFSASFWDCCQKRVTVFQYI 306  
QY 335 GELCRYLLNQPRAENHOVMAAGNLGROSINTWFSRPHIPQVAEYFGATECNSLG 394  
DB 307 GELCRYLVNQPSPKAEFDHKVRLAVGSLRPDTWERFLRRFGPLQILETYGTEGNVATF 366  
QY 395 NFDQVAGCGFNSRLISFYPIRLVRVNEDTMELIRPDGVCIPCQGPFGQLVGRITQK 454  
DB 367 NYTGRQGVGRASWLYKHIFLPSLIRYDVTMTEGPIRNAQHCMTTSPGEGLLVAPVSOQ 426  
QY 455 DPLRFRDGYLNOGANNK-KIARDVFKKGDQAYLTGDDVLMVMDLGYLYFRDRTGTPRWK 513  
DB 427 SP---FLGYAGAPELAKLKKLVFWSGDVFFNTGDLVLCDEQGFLEHFRDGTIRWK 483  
QY 514 ENVSTVEGTLRLLDMADVAVGVVEVPGTEGAGMAVA-SPTGNCDLERFAQVLEKE 572  
DB 484 ENWATEVAEVLDTLDFLQEVNIYGVTVPGHEGAGMAALRPPQALNLVQLYSHVSEN 543  
QY 573 LPLARPIFLRLPLHKTGYTKFQKTELKKEGDPDAIVKDPFLFYLDAAQGRVVPDQEA 632  
DB 544 LPYPARPLRLOESLATTETFKQKVRMANEGDPDPSLDPLVLDQDIDGAYLPLTPAR 603  
QY 633 YSRIQAGEEL 643  
DB 604 YSALLSGDLRI 614

RESULT 10  
O70550

ID O70550 PRELIMINARY; PRT; 620 AA.  
AC O70550;  
DT 01-AUG-1998 (TEMBLrel. 07, Created)  
DT 01-AUG-1998 (TEMBLrel. 07, Last sequence update)  
DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)  
DE VERY-LONG-CHAIN ACYL-COA SYNTHETASE.  
GN VLACS.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALE/C; TISSUE=WHOLE LIVER;  
RX MEDLINE=98218572; PubMed=9559670;  
RA Berger J., Truppe C., Neumann H., Forss-Petter S.;  
RT "CDNA cloning and mRNA distribution of a mouse very long-chain acyl-  
COA synthetase";  
RL FEBS Lett. 425:305-309(1998).  
DR EMBL; AJ223958; CAA11687.1; -  
DR InterPro: IPR000873; -  
DR Pfam: PF00501; AMP-binding; 1.  
DR PROSITE; PS00455; AMP\_BINDING; 1.  
SQ SEQUENCE 620 AA; 70394 MW; FC59270C41784364 CRC64;

Query Match 30.2%; Score 1021; DB 11; Length 620;  
Best Local Similarity 37.7%; Pred. No. 5.1e-77;  
Matches 235; Conservative 108; Mismatches 248; Indels 32; Gaps 12;

QY 32 LYLGGGWRFIRVFKT-----IRRDFGLVLLKVKARQCLQER--RTVPILFASTVR 85  
DB 5 LYTGLAGLLLLPULLTCCPCYLLQDVRYFLRLANMARVRSYRQRRPVITIRAFLEQAR 64  
QY 86 RHPDXTALFECDTHTWFRQLDEYSSVANFLQAR-GLASGDVAALFMENRFEVLWL 144  
DB 65 KTHPFPFLFR--DETLTAQVDRRSQVARALHDLQGLRGDCVALFMGNEPAYVWIL 122  
QY 145 GMAKLGVEAALINTNLRRDALLHCLTTSARALYFGSEMAAICEVHASL--DPSLSLFC 202  
DB 123 GLLKGLCPMACLINIRAKSLHLCFCQCGAKVLLASPDQAEVLEALPTLKKDASVVFY 182  
QY 203 SCSWEPGAVPSTEHLDPL-LKDPKHLPCDCKGFTKLFYIYTSGLTPKAAIIVVHS 261  
DB 183 SPTSNTNGVDITLDRKDVGSAPETPESWRS--EVTFTTTPAVIYTSGLTPKAAIINH 240  
QY 262 RYRMAALVYGFRA-----RPNDIVDCLPLYSAGNIVIGQCCLHGMVTVIRKFS 315  
DB 241 R-----LWYGTGLAMSSGITAQDVIYTMPLYSALMIGLHGCIVVGATLALRSKS 293  
QY 316 ASRFWDDCIKNTVQYIGELCRYLLNQPRAENHOVMAAGNLGROSINTWFSR 375  
DB 294 ASQFWDCCRYNVTVIQYIGELLYLNTPOKPNDRDHKKVKKALGNLGRDGVWREFIKRF 353  
QY 376 HIPQVAEYFGATECNSLGNFDSQVAGCGFNSRLISFYPIRLVRVNEDTMELIRPDGV 435  
DB 354 GDHIVYEYFTEGNIQFNVPYPRKIGAVGRANLYLQKRVARVELIKYDVDEKDEPVDANGY 413  
QY 436 CIPCQGPFGQLVGRITQKDLRPFQGLY-LNQGANNKIAKDKVKKGDQAYLTGDDVLMV 494  
DB 414 CLKVPKGVGLVCKLTQLTLP---FIGYAGGKTQTEKKLDRVFKKGDIIYFNSGDLIMD 470  
QY 495 ELGYLYFRDRTGDTFRKGENVSTVEGTLRLLDMADVAVGVVEVPGTEGAGMAVA- 553  
DB 471 RENFVYFHDVGRDTRFKGENVATEVADIVGLVDFVEEVNIVGVPPVPGHEGRIGMASLK 530  
QY 554 ASPTGNCDLERFAQVLEKELPLXARPFLRLPELHKTGYTKFQKTELKKEGDPDAIVK 613  
DB 531 IKENYEFNGKKLFQHIAEYLPYARPRFLRIQDTIETGTGTFKKRKYTLMEEGFNPIVD 590  
QY 614 PLFYLDAAQGRVVPDQEAYSRI 636  
DB 591 TLFMDAETTFVDMTENIYNAI 613

RESULT 11  
Q9Y2P5 PRELIMINARY; PRT; 690 AA.  
AC Q9Y2P5; 01-NOV-1999 (TREMELrel. 12, Created)  
DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)  
DT 01-MAR-2001 (TREMELrel. 16, Last annotation update)  
DE VERY LONG-CHAIN ACYL-COA SYNTHETASE HOMOLOG 2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Steinberg S.J., Wang S.J., Watkins P.A.;  
RT "Human Very Long-Chain Acyl-CoA Synthetase Homolog 2.";  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF064255; AD29444.1; -  
DR InterPro: IPR000873;  
DR Pfam: PF00501; AMP-binding; 1.  
DR PROSITE: PS00455; AMP-BINDING; 1.  
SQ SEQUENCE 690 AA; 75384 MW; 011313424D794546 CRC64;

Query Match 29.1%; Score 984.5; DB 4; Length 690;  
Best Local Similarity 37.7%; Pred. No. 7e-74; Indels 41; Gaps 15;  
Matches 252; Conservative 93; Mismatches 283;

QY 2 LLGASLVGVLLSKVLK---PWTQVGFSLFLYLG-----SGGWRFRVFIKTIR 50  
DB 36 LLGPTCCVLLGLAMLARPMWPGVPHGLSLAAALATLLPARLPGLRWLPA----- 89  
QY 51 RDIFGVLVLKAKVRQCL--QERRFVPLPASTVRRHPDKTALIFEGTDT-HWTFROL 107  
DB 90 -DVIFLAKILHLGKIRGLSROPPTDFVDAFERRARAQPGALLVMTGPGAGSVTFGEL 148  
QY 108 DEYSSSVANFLOAR-----GLASDVAAIFMENRNEFVGL--MLGMKLGVEAALINTNL 160  
DB 149 DARACQAAWALKAEGLDPASLCAGEPTALLVLAQVAPALCMWLGAKLGCPTAWINPHG 208  
QY 161 RRDALLHCLTTSRRARLVFGSEMAICEVHASLD-PSLSLFCSSGSGWEPGAVPPSTEHL 219  
DB 209 RGMPLAHSLSSGARVLVWDPDLRESLEILPKLOAENIRCFYLSHTSP---TPGVGALG 265  
QY 220 PLLKDAKPH-LPSCPDKGFT--DKLEFVYTSGTGLPKAAIVVHSRYRMAALVYGF 276  
DB 266 AALDAAPSHVPADLRAGITWRSPALFYTSGTGLPKPAILLHVRVLMKMLSL-GA 324  
QY 277 RPNDIVDCLPLXHSAGNIYGICQLLHGMTVVRKKSASRFWDDCIKYNCTIVYIGE 336  
DB 325 TADDVYTVLPLXHVMLGVVILGCLDLGATCVLAPKFTSCFWDCCRQHGVTVILYVGE 384  
QY 337 LCRYLLNPPREAEHQVMAINGLNGLRQSIWTFNSSRFHLPQVAFYEGATECNSIGNF 396  
DB 385 LLRYLCNIPQOPEDRTHVRLAMNGLRADVWETFOQRFPIRIWEYVTEGMMGLVNY 444  
QY 397 DSQVACGFNSRLSFVYPIRLVRNEDTMELIRGPDGVCIPQGPPEGLVGRITQKDP 456  
DB 445 VRCGALGKNSCLRLMSPFELVQFDMEAEPVRDNGFCIPVGLGEPGLLTKVVSQQP 504  
QY 457 LRREFDGYN-OGANNKIAKDVFKKQOAYITGDVLVMDLGLYIFRDRGTDFRWKGEN 515  
DB 505 ---FVGVRGPRELSERKLVNRVSGDVTYNTGDVLAMDREGEFLYFRDLRGDTFRWKGEN 561  
QY 516 VSTTEVECTLSRLDMADVAVYGEVPGTEGRAGMAAV-ASPTGNCDLERFAQVLEKELP 574  
DB 562 VSTEVEGVLSDQVDFLOQNVNIGVCVPCGCKGVMAAVQLAPGOTTFGEKLYQHVRAWLP 621  
QY 575 LYARPIFLRLPELHKTGTYYKFKTELKKEGFDPAIVKDPFLFYLDLDAOKGRVYLDQAYS 634  
DB 622 AYATPHFIRIQADMEVTSFTKLMKTRLVREGFNVGIVVDPLFVLNDRAQSFRLTAEMQ 681

QY 635 RIQAGEEKL 643  
DB 682 AVCEGTWRL 690

RESULT 12  
Q9ES38 PRELIMINARY; PRT; 690 AA.  
AC Q9ES38; 01-MAR-2001 (TREMELrel. 16, Created)  
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMELrel. 16, Last annotation update)  
DE BILE ACID COA LIGASE.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=SPRAGUE-DAWLEY;  
RC Falany C.N., Xie X., Wheeler J., Wang J., Barnes S.;  
RT "Molecular cloning and expression of rat liver bile acid CoA ligase.";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF242189; AAG09770.1; -  
KW Ligase.  
SQ SEQUENCE 690 AA; 76265 MW; D5BE8317758DA59 CRC64;

Query Match 28.6%; Score 969.5; DB 11; Length 690;  
Best Local Similarity 37.8%; Pred. No. 1.3e-72; Indels 29; Gaps 12;  
Matches 233; Conservative 106; Mismatches 249;

QY 46 IKTIRDFGLVLLKVKAKVRQCLQERRTPVILFASFVRR---HPDKTALIFEGTD-T 100  
DB 84 LRLWHLKDVAFKLLFYGLNLRRL--NRHPPELFDALEQQAQARPDQVALVCTGSEGC 141  
QY 101 HWTFQRLDEYSSSVANFLOARGLAS-----GDVAALFMENRN-EFVGLWLGMAKLGVEA 153  
DB 142 SITNELNAKACAAWALKAKLKEATIQEDKGATAILVLPKSKISALSIVFGLAKLGPV 201  
QY 154 ALINLNRDALLHCLTTSRRARLVFGSEMAICEVHAS-LDPSLSLFCSSGSGWEPGAVP 212  
DB 202 AWINPHSGMPLHSHVQSSGASVLVDPDLQENLEEVLPKLAENIRCFYLGHSPP---T 258  
QY 213 PTEHLDPKLLKADAPKHLPCSDPK-----GFTDKLFYIYTSCTTGLPKRAIVVHSRYRMA 267  
DB 259 PGVEALGAALDAAPS--DPVPAKLRAKIKKSPAFIYTSCTTGLPKPAILSHERVIOMS 316  
QY 268 ALVYVGFMRPNDIVDCLPLXHSAGNIYGICQLLHGMTVVRKKSASRFWDDCIKYN 327  
DB 317 NVLSFCGR-TADDVYTVLPLXHVMLGVVILGCLDLGATCVLAPKFTSASRYWAEQYS 375  
QY 328 CTIVQYIGELCRYLLNQPPREAEHQVMAINGLNGLRQSIWTFNSSRPHIPQVAFYCAT 387  
DB 376 VTIVVLYVGEVLYLCNVPGQPEDKKHTRVFALNGLRADVWENFQORFGPIQIWEIYGT 435  
QY 388 EGNCSLGNFDSQVACGFNSRLSFVYPIRLVRNEDTMELIRGPDGVCIPQGPCEGOL 447  
DB 436 EGNVGLMNYVGHGAVGKTSCTFIRMLTFLFVQFDIETAEPRDKQKQFCIPVETGKPGLL 495  
QY 448 VGRITQKDLRRFDGYNLQGNANKKIAKDVFKKQOAYITGDVLVMDLGLYIFRDRGTGD 507  
DB 496 LHKIRKNQPFYLRG--SQDETFRKLVRANVQVGLDYNTGDVLALDQEGFFYFRDLGD 553  
QY 508 TFRWGENVSTTEVECTLSRLDMADVAVYGEVPGTEGRAGMAAV-ASPTGNCDLERFA 566  
DB 554 TFRWGENVSTREVEGVLSILDFLEEVNIVGYTVPCGCKGVMAAVKLAPKOTFGDQKLY 613  
QY 567 QVLEKELPLYPARPIFLRLPELHKTGTYYKFKTELKKEGFDPAIVKDPFLFYLDLDAOKGRV 626  
DB 614 QHVRSLPAYATPHFIRIQDLSLEITNTYKLVKSQALAREGFDVGVIADPLYLIDNKAETFR 673





